



KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and
Applied Economics

Department of Economics

Is it fair to "make work pay"??

by

Roland Iwan LUTTENS
Erwin OOGHE

Public Economics

Center for Economic Studies
Discussions Paper Series (DPS) 04.10
<http://www.econ.kuleuven.be/ces/discussionpapers/default.htm>



**DISCUSSION
PAPER**

Is it fair to “make work pay”?*

Roland Iwan Luttens[†] and Erwin Ooghe[‡]

December, 2004.

Abstract

The design of the income transfer program for the lower incomes is a hot issue in current public policy debate. Should we stick to a generous welfare state with a sizeable basic income, but high marginal tax rates for the lower incomes and thus little incentives to work? Or, should we “make work pay” by subsidizing the work of low earners, but possibly at the cost of a smaller safety net? We think it is difficult to answer this question without making clear what individuals are (held) responsible for and what not. First, we present a new fair allocation, coined a Pareto Efficient and Shared resources Equivalent allocation (PESE), which compensates for different productive skills, but not for different tastes for working. We also characterize a fair social ordering, which rationalizes the PESE allocation. Second, we illustrate the optimal second-best allocation in a discrete Stiglitz (1982, 1987) economy. The question whether we should have regressive or progressive taxes for the low earners crucially depends on whether the low-skilled have a strictly positive or zero skill. Third, we simulate fair taxes for a sample of Belgian singles. Our simulation results suggest that “making work pay” policies can be optimal, according to our fairness criterion, but only in the unreasonable case in which most of the unemployed are not willing to work.

1 Motivation

Focussing on the tax-benefit system as a whole, many European countries combine a sizeable basic income with high marginal taxes for the low income earners. These programs are praised for their redistributional appeal, directing the largest possible transfer towards the poorest in society. But, at the same time, critics have held these schemes responsible for large unemployment traps, because they do not provide incentives to (start) work(ing). Therefore, some continental European countries —such as Belgium, Finland, France, Germany, Italy and the Netherlands— have proposed and/or introduced tax credit schemes recently, to subsidize the low income earners; see Bernardi and Profeta (2004) for an overview. At the same time, the US and the UK, with a

*We would like to thank François Maniquet, Kristian Orsini and Dirk Van de gaer for useful comments, and Kristof Bosmans and Dirk Hoorelbeke for their euros.

[†]SHERPPA, Ghent University, Belgium, e-mail: roland.luttens@ugent.be. Financial support from the Federal Public Planning Service Science Policy, Interuniversity Attraction Poles Program – Belgian Science Policy (contract no. P5/21) is gratefully acknowledged.

[‡]Postdoctoral Fellow of the Fund for Scientific Research - Flanders. Center for Economic Studies, K.U.Leuven, Belgium. e-mail: erwin.ooghe@econ.kuleuven.ac.be.

much longer tradition in tax credit schemes, have reinforced the role of their tax credits. The increased policy interest for such “making work pay” schemes is its ability to tackle two problems at the same time. It has a positive effect on employment (the number of people working and, to a lesser extent, the aggregate labour hours), while it increases the income of poor households; see Pearson and Scarpetta (2000) for an overview.

While “making work pay” schemes may attain desirable objectives, it is not clear whether it is also optimal to “make work pay” for a given budget constraint. The “welfarist” optimal income tax literature consists of three canonical models, depending on whether labour supply responses are modelled intensively and/or extensively (Heckman, 1993). *First*, in a Mirrlees (1971) economy, individuals respond via the intensive margin, i.e., by varying their labour hours or effort. Marginal taxes should be non-negative everywhere (Mirrlees, 1971), which excludes the possibility of subsidizing work. At the bottom, the marginal tax has to be zero, but only in case everybody works (Seade, 1977). Once there exists an atom of non-workers, the marginal tax rate has to be positive (Ebert, 1992) and, according to some numerical simulations (Tuomala, 1990), rather high. Using the empirical earnings distribution, high and decreasing marginal taxes at the bottom seem to survive (Diamond, 1998, Kanbur and Tuomala, 1994, Piketty, 1997 and Saez, 2001). To conclude, in a Mirrlees economy, high marginal tax rates at the bottom seem optimal. *Second*, in Diamond’s (1980) approach, individuals respond via the extensive margin, i.e., they choose to work or not. Marginal tax rates can be negative, suggesting at least the possibility of subsidizing the work of low earners. *Third*, Saez (2002) presents a unifying framework where individuals can respond via both margins. Support for one of both income transfer schemes depends on the relative importance of both response margins and on the redistributive tastes of government. Saez’ benchmark simulation suggests a sizeable basic income (around \$7300/year), but, combined with a tax exemption at the bottom (for incomes up to \$5000/year).

In the same year of Mirrlees’ (1971) seminal contribution, Rawls (1971) criticizes the welfarist approach. In the aftermath of Rawls’ influential work, many alternative theories of distributive justice were proposed. Although very diverse in equalisandum, they almost all have Dworkin’s (1981) cut in common. Dworkin claims that not all individual characteristics can (should) be considered as morally arbitrary. Therefore one has to make a clear cut between endowments and ambitions. He introduces personal responsibility: individuals are responsible for their ambitions —as long as they identify with them— but not for their endowments. As a consequence, a fair distribution scheme should be ambition-sensitive, but endowment-insensitive.

In an optimal income tax setting, fairness could require to compensate for differences in productive skill (endowment), but not for differences in taste for working (ambition). Schokkaert *et al.* (2004) introduce such fairness considerations in different ways and calculate the corresponding optimal linear income tax, which turns out to be positive.¹ Allowing for non-linear tax schemes, results change drastically. Boadway *et al.* (2002) analyze the optimal non-linear income tax according to a weighted utilitarian or maximin social planner where different weights are chosen for different tastes. Fleurbaey and Maniquet (2002) characterize a fair social ordering to analyze non-linear income taxes. In both studies, negative marginal taxes (for the low income earners) may be optimal.

¹Roemer *et al.* (2001) consider the education level of the parents as the compensating variable.

In the next section we present a new fair allocation, coined a Pareto Efficient and Shared resources Equivalent (PESE) allocation. As the name suggests, the optimal allocation is Pareto efficient and all individuals are indifferent between their bundle and what they would get if it were physically possible to divide or share all resources, including the productive skills. Section 3 characterizes a fair social ordering, which rationalizes the PESE allocation. In section 4, we introduce a “discrete” Stiglitz (1982, 1987) economy with (i) four types of individuals (defined by a low or high productive skill and a low or high taste for working) and (ii) a government who wants to install fair taxes, but cannot observe individuals’ type. We show that fairness recommends regressive taxation for the low earners, as long as the low-skilled individuals have a strictly positive skill. In case the low-skilled have a zero skill, only progressive taxes can be optimal. In section 5, we simulate fair taxes for Belgian singles, while carefully paying attention to the calibration of the compensation (hourly wages) and responsibility (taste for working) variable. Our simulation results suggest that “making work pay” policies can be optimal—according to our fairness criterion—but only in the unreasonable case in which most of the unemployed are not willing to work.

2 Equality of resources revisited

When all resources in society are alienable and divisible, Dworkin proposes to divide resources equally (endowment insensitivity), followed by an auction to reallocate resources according to taste (ambition sensitivity). This leads to a Pareto efficient and envy-free allocation. To study fair income taxation, however, we have to introduce productive resources (skills), which are not alienable, and therefore a problem arises. In production economies, Pareto efficient and envy-free allocations do not exist, in general.

A first class of solutions tries to extend the above Dworkinian auction by assigning property rights over leisure. Varian (1974) analyzes two, rather extreme, solutions. One may divide consumption goods equally and either (i) assign each individual his own leisure, or (ii) give each individual an equal share in each of the agents’ (including his own) leisure time. After trade, the resulting competitive (and hence Pareto efficient) equilibria are called respectively (i) wealth-fair and (ii) income-fair. In the wealth-fair allocation, productive talents are a private good and the resulting allocation does not compensate at all for inabilities. In the income-fair allocation, productive talents are a public good. The high-skilled has to buy back his expensive leisure and is therefore punished for being a high skill type, resulting in a slavery of the talented. Intermediate solutions exist where skills are neither purely private, nor purely public (Fleurbaey and Maniquet, 1996, Maniquet, 1998).

A second class of solutions starts from the concept of fair-equivalence. Pazner and Schmeidler (1978) define an allocation to be fair-equivalent if everyone is indifferent between his bundle in this allocation and the bundle he would receive in a “hypothetical” fair, i.e., envy-free, allocation. It then suffices to define an interesting “hypothetical” fair allocation and to look whether there exist Pareto efficient ones, among all fair-equivalent allocations. The resulting allocation is called a Pareto efficient and fair-equivalent (PEFE) allocation.

Pazner and Schmeidler (1978) propose an egalitarian allocation—an allocation where everybody consumes the same consumption-leisure bundle—as the fair one, which leads

to Pareto efficient and egalitarian-equivalent (PEEE) allocations. We propose a different fair allocation, which we coin a “shared resources” allocation. This is the allocation which would result, if it were (physically) possible to divide or share all resources, including the productive ones. To make this idea more precise, we introduce some notation.

A fixed number of individuals, denoted $i \in N = \{1, \dots, n\}$, differ in skills and preferences. Skill $s \in \mathbb{R}_+$ defines production (called gross income in the sequel) in a linear way, or $y = s\ell$, with $\ell \in [0, 1]$ the amount of labour. We denote a skill profile by $\mathbf{s} = (s_1, \dots, s_n) \in \mathbb{R}_+^n$. Taste for working is represented by a continuously differentiable utility function

$$U : \mathbb{R} \times [0, 1] \rightarrow \mathbb{R} : (c, \ell) \mapsto U(c, \ell),$$

which is strictly increasing (resp. strictly decreasing) in consumption c (resp. labour ℓ) and strictly quasi-concave. We call \mathcal{U} the corresponding set of utility functions and, normalizing the consumption price equal to one, we refer to c as the net income in the sequel. A utility profile is denoted by $\mathbf{U} = (U_1, \dots, U_n) \in \mathcal{U}^n$. An economy $e = (\mathbf{s}, \mathbf{U})$ is completely defined by a skill and a utility profile; all economies are gathered in a set $\mathcal{E} = \mathbb{R}_+^n \times \mathcal{U}^n$.

Individuals are (held) responsible for their tastes, but not for their skills. Therefore, we want to compensate individuals for different outcomes which are only due to different skills, but not for different outcomes which are only caused by different tastes for working. In case skills are alienable —think, e.g., of individuals as farmers who receive, as a matter of brute bad luck, either a blunt or a whetted scythe (the skill s) to harvest crops (the consumption c)— there is a particularly simple and attractive way to obtain a fair allocation:

- (a). each individual pays (or receives) the same lump-sum amount of money,
- (b). each individual can use each skill (including his own) for a time equal to $\frac{1}{n}$ at most.

As such, all individuals would end up with the same opportunity set. In the sequel, we call this opportunity set the “shared resources” opportunity set and the resulting allocation (which ultimately depend on the tastes in society) is called the “shared resources” allocation.

To illustrate these concepts, suppose (i) there are only two skill types possible in society, say low (L) and high (H), which are equally represented in the skill pool \mathbf{s} , and (ii) there are only two tastes for working possible, also called low (L) and high (H). An allocation $\mathbf{z} = (z_{LL}, z_{LH}, z_{HL}, z_{HH})$ contains one bundle $z_{st} = (c_{st}, \ell_{st})$ for each of the four types st , with s referring to the skill (low or high) and t referring to the taste (low or high). Figure 1 illustrates the opportunity sets and (resulting) allocations in case (a) each individual receives the same lump-sum amount of money a , but productive resources are not shared, and (a)+(b) each individuals receives the same lump-sum amount a and also the productive resources are shared (each individual can work with each of the skills half-time at most).

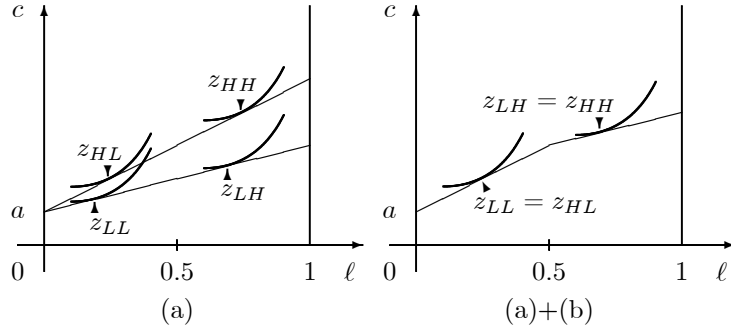


Figure 1: Opportunity set/allocation change when sharing productive resources.

Sharing productive resources is not technically feasible in many cases. Labour market productivities, due to inborn characteristics such as intelligence, talents, handicaps and so on, are typically inalienable. Still, we could consider the allocation, which would arise if it were possible to divide and share all resources equally, as an interesting “hypothetical” case. However, the resulting hypothetical “shared resources” allocation is not Pareto efficient, in general. Therefore, we propose to focus on Pareto Efficient and Shared resources Equivalent (PESE) allocations. Figure 2 illustrates a PESE allocation for an economy defined by the same assumptions (i) and (ii) as in figure 1. Due to Pareto efficiency, the slope of an individual’s indifference curve must be equal to his skill level in his bundle. Given the skill and preference technology, there exists a unique PESE allocation for each value of a .

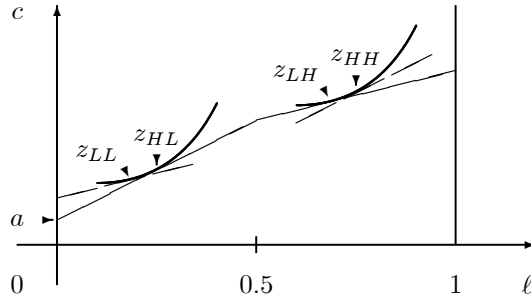


Figure 2: A Pareto efficient and shared resources equivalent allocation.

3 A “shared resources” social ordering

In case it is possible to recognize the less from the more productive and the lazy from the hard-working individuals, we can choose among all Pareto efficient and “shared resources” equivalent allocations described in the previous section. However, it is not always possible to observe types. To proceed in such a second-best setting, it is more convenient to characterize a corresponding “shared resources” social ordering.

We define our well-being concept, which is closely linked to the PESE allocation. Let $\mathcal{Z} = (\mathbb{R} \times [0, 1])^n$ be the set of allocations $\mathbf{z} = (z_i)_{i \in N}$, containing one bundle $z_i = (c_i, \ell_i)$ for each individual i in N . We get (an explanation follows):

WELL-BEING: For each allocation $\mathbf{z} \in \mathcal{Z}$, the vector of well-being levels $\mathbf{w} = (w_i)_{i \in N} \in \mathbb{R}^n$ is defined by the amounts of money w_i which would make individual i indifferent between (i) receiving (or paying) this amount of money w_i and sharing all productive resources equally (in time), and (ii) his actual bundle x_i . Because the well-being vector \mathbf{w} depends on the allocation \mathbf{z} and the economy $e = (\mathbf{s}, \mathbf{U})$, we write $\mathbf{w} = \mathbf{W}(\mathbf{z}, e)$, with $w_i = \mathbf{W}_i(\mathbf{z}, e)$.

A few observations need to be stressed here. *First*, for the PESE allocation presented in figure 2, the well-being levels are the same for all individuals and equal to a . More general, in any “shared resources” equivalent allocation, individuals end up with the same well-being and, vice versa, if all individuals have the same well-being in an allocation, the latter must be “shared resources” equivalent. *Second*, well-being has to be interpreted as a “relative” measure of fair treatment, in the spirit of the PESE allocation. If two individuals have the same well-being, they have been treated equally fair, because both individuals are indifferent between their actual bundle and the bundle they would choose if (a) they receive the same lump-sum amount of money and (b) all productive resources are shared equally. If one individual has a strictly lower well-being compared to another, he has been treated unfair with respect to the other, because both individuals are indifferent between their actual bundle and the bundle they would choose if (b) all productive resources are shared equally, but (a) the former individual receives a strictly lower lump-sum amount of money. *Third*, higher utility (and thus a higher indifference curve) also leads to a higher well-being level. As such, our definition of well-being corresponds with one specific, but, according to us, interesting cardinalization of the utility functions.

A rule f maps economies into orderings, or $f : \mathcal{E} \rightarrow \mathcal{R} : e \mapsto R_e = f(e)$, with \mathcal{R} the set of all orderings (complete and transitive binary relations) defined over allocations \mathbf{z} in \mathcal{Z} ; call P_e and I_e the corresponding asymmetric and symmetric relation. We define some properties for f . Our Pareto principle is equal to Pareto Indifference and the Weak Pareto principle together, i.e., if everyone is indifferent between allocations \mathbf{z} and \mathbf{z}' , then \mathbf{z} should also be socially indifferent to \mathbf{z}' and if everyone strictly prefers allocation \mathbf{z} to \mathbf{z}' , then \mathbf{z} should also be socially strictly preferred to \mathbf{z}' . Anonymity requires that the names of the individuals do not matter. Formally:

PARETO: For each economy $e \in \mathcal{E}$ and for all allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$: If $U_i(z_i) = U_i(z'_i)$ for all $i \in N$, then $\mathbf{z} I_e \mathbf{z}'$. If, $U_i(z_i) > U_i(z'_i)$ for all $i \in N$, then $\mathbf{z} P_e \mathbf{z}'$.

ANONIMITY: For each economy $e \in \mathcal{E}$, for each allocation $\mathbf{z} \in \mathcal{Z}$ and for each permutation $\pi : N \rightarrow N$ over individuals: If $U_i = U_j$ for all $i, j \in N$, then $\mathbf{z} I_e \pi(\mathbf{z})$, with $\pi(\mathbf{z}) = (z_{\pi(1)}, \dots, z_{\pi(n)})$.

In line with the idea to compensate for differences in outcomes which are only due to differences in skills, compensation (Fleurbaey and Maniquet, 2002) requires that a Pigou-Dalton transfer (in terms of net income) from a rich to a poor individual with the same preferences and the same labour should be welfare improving:

COMPENSATION: For each economy $e \in \mathcal{E}$, for all allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$ and for all individuals $i, j \in N$: If (i) $\ell_i = \ell_j$, $\ell'_i = \ell'_j$ and $U_i = U_j$, (ii) $\exists \delta > 0$ such that $c'_i = c_i + \delta < c_j - \delta = c'_j$ and (iii) $z_k = z'_k$ for all $k \neq i, j$, then $\mathbf{z} R \mathbf{z}'$.

Finally, in line with (i) our well-being definition and (ii) the idea that individuals are responsible for their tastes, Well-being Independence requires the ranking of two allocations to be the same (i) whenever they give rise to the same well-being vector, (ii) irrespective of the utility profile:

WELL-BEING INDEPENDENCE: For all economies $e = (\mathbf{s}, \mathbf{U})$, $e' = (\mathbf{s}, \mathbf{U}') \in \mathcal{E}$ and for all allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$: If $\mathbf{W}(\mathbf{z}, e) = \mathbf{W}(\mathbf{z}, e')$ and $\mathbf{W}(\mathbf{z}', e) = \mathbf{W}(\mathbf{z}', e')$, then $\mathbf{z} R_e \mathbf{z}' \Leftrightarrow \mathbf{z} R_{e'} \mathbf{z}'$.

Given these axioms, we should focus on the minimal well-being in society, or:²

Proposition 1. *If a rule $f : \mathcal{E} \rightarrow \mathcal{R}$ satisfies Pareto, Anonymity, Compensation and Well-being Independence, then, for each economy $e \in \mathcal{E}$ and for all allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$: $\min \mathbf{W}(\mathbf{z}, e) > \min \mathbf{W}(\mathbf{z}', e)$ implies $\mathbf{z} P_e \mathbf{z}'$.*

The “shared resources” social ordering has some formal similarity with Fleurbaey and Maniquet’s (2000) \tilde{s} -implicit budget leximin function, where \tilde{s} is a reference skill level. In our case, the reference skill \tilde{s} is piece-wise linear and endogenously defined by the skill pool in society. As such, laissez-faire allocations are selected in case all individuals have the same skill.

4 Fair taxes: theory

In the previous section, we characterize a fair social ordering, inspired by the PESE allocation. In this section, we analyze what happens when the government uses this fair social ordering to calculate optimal taxes in a discrete Stiglitz economy (1982, 1987) with four types, which are not observable to the government.

All individuals in $N = \{1, \dots, n\}$ can have four types, denoted $(s, t) \in S \times T$, where s is the skill level and t the taste for working; we abbreviate types as $st \in ST$. Each type st is represented by $n_{st} > 0$ individuals, with $n = \sum_{st \in ST} n_{st}$. Skills can be low or high, or $s \in S = \{L, H\}$, with $0 < L < H$; later on, we come back to the issue of zero skills. Tastes for working can also be low or high, or $t \in \{L, H\}$, which correspond with a utility function U_t .³

As before, utility functions belong to \mathcal{U} , but we impose some additional properties. Let V_{st} represent the preferences in the consumption-income space for type st , more precisely $V_{st} : \mathbb{R} \times [0, s] \rightarrow \mathbb{R} : (c, y) \mapsto V_{st}(c, y) \equiv U_t(c, \frac{y}{s})$.⁴ We impose two additional

²All proofs can be found in appendix A.

³The exact scalars do not matter here, so we stick to the notation of L to denote low taste for working and/or low-skilled and $H > L$ to denote high taste for working and/or high-skilled.

⁴Whenever $s = 0$, we define $V_{st}(c, y) = c$ for all $(c, y) \in \mathbb{R} \times \{0\}$.

properties on the utility functions U_t ; see Stiglitz (1982, 1987) for the first and Boadway *et al.* (2002) for the second property:

SINGLE-CROSSINGNESS: A higher taste for working t corresponds with a lower marginal rate of substitution (denoted $MRS_t = -\frac{\partial U_t/\partial \ell}{\partial U_t/\partial c}$), expressing the view that individuals with a higher taste for working require less compensation (in terms of net income c) to work a little bit longer. Formally: $MRS_L > MRS_H$ in $\mathbb{R} \times [0, 1]$.

INDISTINGUISHABLE MIDDLE TYPE: The types LH and HL have the same preferences in the consumption-income space. Formally, there exists a continuous and strictly increasing function $\phi : \mathbb{R} \rightarrow \mathbb{R}$, such that $V_{LH} = \phi \circ V_{HL}$ in $\mathbb{R} \times [0, L]$.

Both assumptions together, the marginal rates of substitution in consumption-income space (denoted $MRSY_{st} = -\frac{\partial V_{st}/\partial y}{\partial V_{st}/\partial c}$) are also single-crossing, more precisely:

$$\begin{aligned} MRSY_{LL} &> MRSY_{LH} = MRSY_{HL} > MRSY_{HH}, \text{ in } \mathbb{R} \times [0, L] \text{ and} \\ MRSY_{HL} &> MRSY_{HH} \text{ in } \mathbb{R} \times [L, H]. \end{aligned}$$

We focus in the sequel on allocations $\mathbf{x} = (x_{LL}, x_{LH}, x_{HL}, x_{HH})$ in consumption-income space, thus $\mathbf{x} \in \mathcal{X} = (\mathbb{R} \times [0, L])^2 \times (\mathbb{R} \times [0, H])^2$, containing one bundle $x_{st} = (c_{st}, y_{st})$ for each type $st \in ST$. The program of the government is to find the best allocation(s) \mathbf{x} —“best” according to the fair social ordering defined in proposition 1— subject to (i) incentive compatibility constraints (no type envies another type’s bundle) and (ii) a feasibility constraint (the sum of all taxes is larger than the government requirement $g \in \mathbb{R}$). Recall our definition of well-being in the previous section; with a slight abuse of notation, we write the well-being of type st in allocation \mathbf{x} as $w_{st} = \mathbf{W}_{st}(\mathbf{x})$. We get:

$$\max_{\mathbf{x} \in \mathcal{X}} \min_{st \in ST} (\mathbf{W}_{st}(\mathbf{x}))_{st \in ST} \text{ subject to} \quad (*)$$

INCENTIVE COMPATIBILITY CONSTRAINTS $IC_{st, (st)'}'$:

$$\begin{aligned} V_{st}(x_{st}) &\geq V_{st}(x_{(st)'}), \forall st \in \{H\} \times T, \forall (st)' \in ST, \\ V_{st}(x_{st}) &\geq V_{st}(x_{(st)'}), \forall st \in \{L\} \times T, \forall (st)' \in ST \text{ with } y_{(st)'} \leq L. \end{aligned}$$

FEASIBILITY CONSTRAINT:

$$\sum_{st \in ST} n_{st} (y_{st} - c_{st}) \geq g.$$

Our first result tells us that the lowest income type, the “undeserving poor” with type LL , must always receive less subsidies (or pay higher taxes) than the second lowest income type, the “hard-working poor” with type LH . This result suggests that it is optimal —according to our fair social ordering— to “make work pay” by subsidizing the low earners:

Proposition 2. Consider a four type economy with skills $0 < L < H$ and tastes represented by utility functions $U_L, U_H \in \mathcal{U}$, which satisfy single-crossingness and indistinguishable middle type. Consider a government who optimizes the program defined by (*). In an optimal allocation $\mathbf{x}^* \in \mathcal{X}$ we must have $y_{LL}^* - c_{LL}^* \geq y_{LH}^* - c_{LH}^*$.

We have to put this result in perspective, however. Although it is reasonable to assume that all individuals (with a capacity for work) have strictly positive productive skills, individuals might be constrained in their choice due to labour market frictions. Minimum wage laws, rationing and so on, may prevent individuals, in particular the low-skilled, from working. Suppose, in our four type economy, that the low-skilled individuals, are willing, but cannot work, due to such constraints, which are beyond their responsibility. In such a case, their skills are nullified. This turns proposition 2 round, or, taxes must be progressive for the bottom incomes:

Proposition 3. Consider a four type economy with skills $L = 0 < H$ and tastes represented by utility functions $U_L, U_H \in \mathcal{U}$, which satisfy single-crossingness and indistinguishable middle type. Consider a government who optimizes the program defined by (*). In an optimal allocation $\mathbf{x}^* \in \mathcal{X}$ we must have $y_{LL}^* - c_{LL}^* = y_{LH}^* - c_{LH}^* \leq y_{HL}^* - c_{HL}^*$.

Both proposition 2 and 3 are based on simple fictitious economies. Furthermore, in proposition 3 we consider a rather extreme case in which all unemployed (the low-skilled) are not able to work. In the next section, we simulate fair taxes for a sample of Belgian singles. It allows us to focus on (i) more realistic economies with many different types and, more importantly, on (ii) different and more realistic scenarios concerning the ability of the unemployed to work. The different scenarios in (ii) have a crucial impact on the tax-benefit scheme for the low earners.

5 Fair taxes: simulation results

5.1 Calibration

We use a sample of singles from the 1997 wave of the Panel Study for Belgian Households; we only include singles with a capacity for work (students, pensioners, sick, or handicapped singles are excluded). We observe (i) the pre-tax yearly labour income y , (ii) the amount of labour ℓ , normalized such that $0 \leq \ell \leq 1$, where $\ell = 1$ corresponds with 2925 hours, i.e., 45 weeks times 65 hours, (iii) the gross hourly wage rate σ (only observed for those who worked, i.e., both $y, \ell \neq 0$) which leads to a gross yearly wage rate $s = 2925\sigma$ and (iv) the total net unemployment benefit β (only observed for those who were partly or completely unemployed in 1997) from which we derive the net yearly unemployment benefit $b = \frac{\beta}{1-\ell}$, i.e., the net unemployment benefit one would obtain if full-time unemployed ($\ell = 0$).

First, we consider all individuals for which we possess all of the above information. We consider quasi-linear preferences (which exclude income effects) represented by utility functions:⁵

$$U_t : \mathbb{R} \times [0, 1] \rightarrow \mathbb{R} : (c, \ell) \mapsto U_t(c, \ell) = c - \frac{1}{t} \frac{\varepsilon}{1 + \varepsilon} \ell^{\frac{1+\varepsilon}{\varepsilon}},$$

⁵Due to quasi-linearity, other non-labour income (e.g., due to rents, gifts, alimony, child allowances) does not matter for the labour choice of an individual. Furthermore, we assume that non-labour income falls within the responsibility of an individual and it is therefore excluded from our analysis.

with t the taste parameter (possibly different for different individuals) and ε the labour supply elasticity (the same for all individuals). Preferences in consumption-income space become:

$$V_{st} : \mathbb{R} \times [0, s] \rightarrow \mathbb{R} : (c, y) \mapsto V_{st}(c, y) = c - \frac{1}{s^{\frac{1+\varepsilon}{\varepsilon}} t} \frac{\varepsilon}{1+\varepsilon} y^{\frac{1+\varepsilon}{\varepsilon}}.$$

The net income of a Belgian single equals $y - \tau_{97}(y) + b(1 - \frac{y}{s})$, with $\tau_{97}(\cdot)$ the actual tax system for singles in Belgium in 1997 (reported in appendix B) and $b(1 - \frac{y}{s})$ the benefit when working $\ell = \frac{y}{s}$ units of time. Both tax and benefit parts separately, as well as the resulting budget set (the solid line), are illustrated in figure 3.

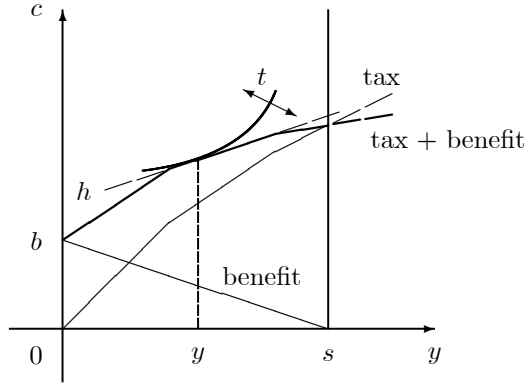


Figure 3: Calibration of the taste parameter t .

We calibrate t such that the choice of y is rationalized for each single. More precisely, the slope of the individual's budget set at y , denoted $h(y)$, should be equal to the marginal rate of substitution between consumption and gross income for the quasi-linear preferences V_{st} ; we get⁶

$$t = \frac{1}{s} \frac{1}{h(y)} \left(\frac{y}{s} \right)^{\frac{1}{\varepsilon}}, \text{ with } h(y) = 1 - \tau'_{97}(y) - \frac{b}{s}.$$

Second, since we could only observe gross yearly wages s (resp. net yearly unemployment benefits b) for individuals who worked in 1997 (resp. individuals who received unemployment benefits in 1997), we complete our dataset by imputing values for s and b , whenever unobserved, via a Heckman selection model. Thus, in estimating s and b , we correct for a possible sample selection bias, due to the fact that we only observe wages s for those who worked and benefits b for those who were (permanently or temporarily) unemployed. The variables used for the imputation as well as the estimation results are described in appendix C.

We end up with a heterogeneous sample of 621 singles who differ in skills s and tastes t , which drive their labour market behaviour; appendix D contains some descriptive statistics for our dataset. Two points are worth mentioning here. First, the non-responsibility parameter s and the responsibility parameter t in our dataset are barely correlated: using a low labour supply elasticity $\varepsilon = 0.1$ for singles, the correlation between s and

⁶Requiring $t > 0$, individuals with $h \leq 0$ were dropped out of the sample (17 observations).

t equals -0.071 , suggesting independently distributed skills and tastes. With a strong correlation, compensating for skills only (and not for tastes) would be a dubious exercise. Second, given the nature of our quasi-linear preferences, all unemployed individuals receive a taste for working $t = 0$. To put it differently, all unemployed are considered unwilling to work. We relax this crucial assumption later on.

5.2 Results

Rather than using allocations as in the government program (*), we use a piece-wise linear tax-benefit scheme as our instrument to approximate a non-linear tax scheme. As we are mainly concerned with the bottom incomes, we consider a piecewise linear tax-benefit scheme up to yearly gross earnings of €20000 in steps of €500 and we use a constant marginal tax rate afterwards. Using either a wider range of piecewise linear taxes (up to €80000 in steps of €500) or a finer grid (up to €20000 in steps of €250) does not change our results for the bottom incomes drastically. Remarkably, using a wider range leads to approximately constant marginal tax rates for incomes above €20000, with the exception of the very high incomes. Given such a tax-benefit scheme, individuals choose their best bundle (according to their tastes and skills) and, therefore, incentive constraints are not necessary anymore. For the feasibility constraint, we use the total government requirement (g) in the actual system, which is (in per capita terms) equal to €3851.

5.2.1 The benchmark simulation

No income effects and a low labour supply elasticity do not seem unrealistic for singles; see, e.g., Blundell and MaCurdy (1999) for an empirical assessment. Using a labour supply elasticity $\varepsilon = 0.1$ as a benchmark (BM), figure 4 depicts the chosen bundles in the consumption-gross income space; the dotted line does not represent an optimal tax-benefit schedule, but is only connecting the chosen consumption-gross income bundles. We also report (i) how the benchmark allocation changes when adding participation constraints (BM+PC), i.e., everyone prefers his bundle rather than not participating and receiving the bundle $(0,0)$, and (ii) the Rawlsian optimal allocation, i.e., the one which maximizes the basic income (RAWLS). Our benchmark simulation (BM) is rather extreme, with a negative (yearly) basic income of €-16067, very strong subsidies as soon as individuals start working (up to gross incomes equal to €3000), a progressive part for gross incomes between €3000 and €9500, a small regressive part again for incomes between €9500 and €11500 and progressive taxes afterwards. Adding participation constraints (BM+PC), we obtain a low (yearly) basic income equal to €518, moderate subsidies fading in around €3000 and fading out around €11500, followed by progressive taxes. The Rawlsian case installs a positive basic income equal to €9363 and high positive marginal tax rates for the bottom incomes.

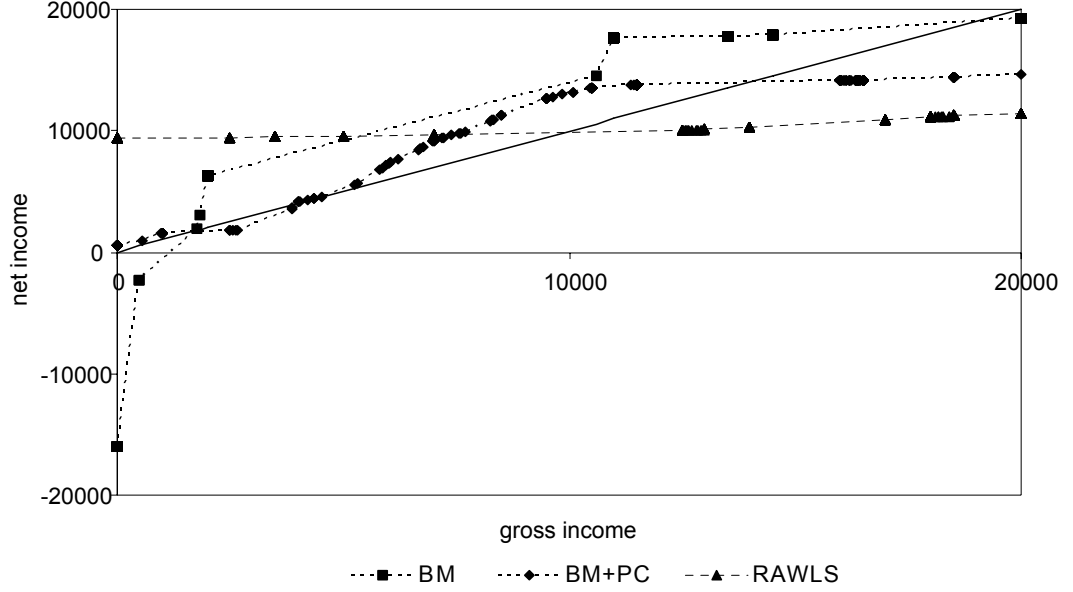


Figure 4: Optimal allocation for the benchmark and the Rawlsian case.

Table 1 summarizes for all schemes and for different groups of individuals G (i) the proportion of individuals $\frac{|G|}{n}$ and (ii) the average tax rate $(\frac{1}{|G|} \sum_{i \in G} (y_i - c_i))$.

	G	$y = 0$	$0 < y \leq 10000$	$10000 < y \leq 20000$	$20000 < y$
BM	proportion	0.19	0.09	0.36	0.36
	avg. tax	16067	-3951	-3765	6955
BM+PC	proportion	0.19	0.11	0.41	0.29
	avg. tax	-518	-1321	561	13104
RAWLS	proportion	0.21	0.14	0.40	0.25
	avg. tax	-9363	-4206	5849	15856

Table 1: Some characteristics for the benchmark and the Rawlsian case.

5.2.2 The impact of ε

In our benchmark simulation, the labour supply elasticity ε equals 0.1. Using a lower $\varepsilon = 0.05$, the correlation between s and t equals -0.049, while for a higher $\varepsilon = 0.2$, the correlation between s and t becomes -0.075. Figure 5 shows the limited impact of varying ε (0.05, 0.1, 0.2) on our optimal allocation. The sequence of regressive and progressive parts turns out to be rather robust to the labour supply elasticity.

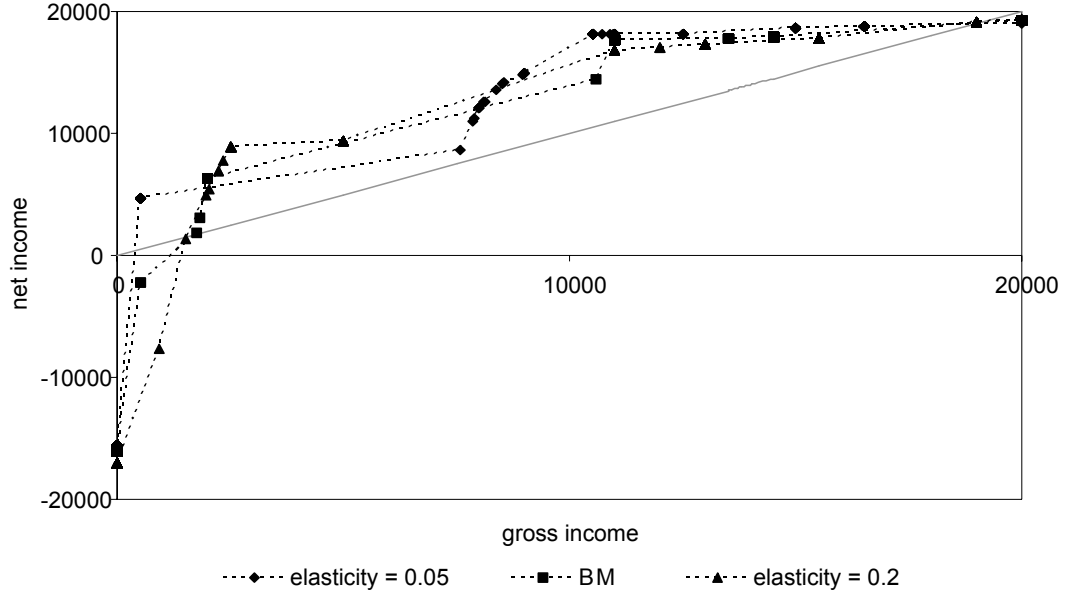


Figure 5: Measuring the impact of varying ε .

5.2.3 What about choice constraints?

As far, individuals who do not work ($y = 0$) are all modelled as individuals who do not want to work ($t = 0$). This is clearly an extreme viewpoint. For example, minimum wage laws in Belgium could keep some individuals (especially those with low skills s) from working and, therefore, our calibration might underestimate their true taste parameter t , thus overestimating their well-being level w . More reasonably, at least some of the observed unemployment must be involuntary, especially in the case of singles. We consider two, rather conservative cases: unemployment is voluntary for the unemployed with either the 90% highest (90%) or the 75% highest (75%) productivities. The results change dramatically.

We proceed as follows. First, we assign (as a start value) to each constrained individual a taste parameter t in the neighbourhood of the average taste parameter of the working and minimize the overall correlation between skills and tastes by deviating from these start values. Afterwards, we keep these individuals constrained at $y = 0$, but use their “true” taste parameter to calculate well-being levels.⁷ Figure 6 presents the benchmark case (BM) —which represents the extreme case where all unemployment is voluntary— together with the 90%- and the 75%-case —where respectively 90% and 75% of the unemployment is voluntary (for those with the higher productivities)— as well as the Rawlsian (RAWLS) case, which maximizes the basic income.

⁷As we do not know whether it is technically possible to alleviate these constraints, we stick to tax instruments to maximize the minimal well-being.

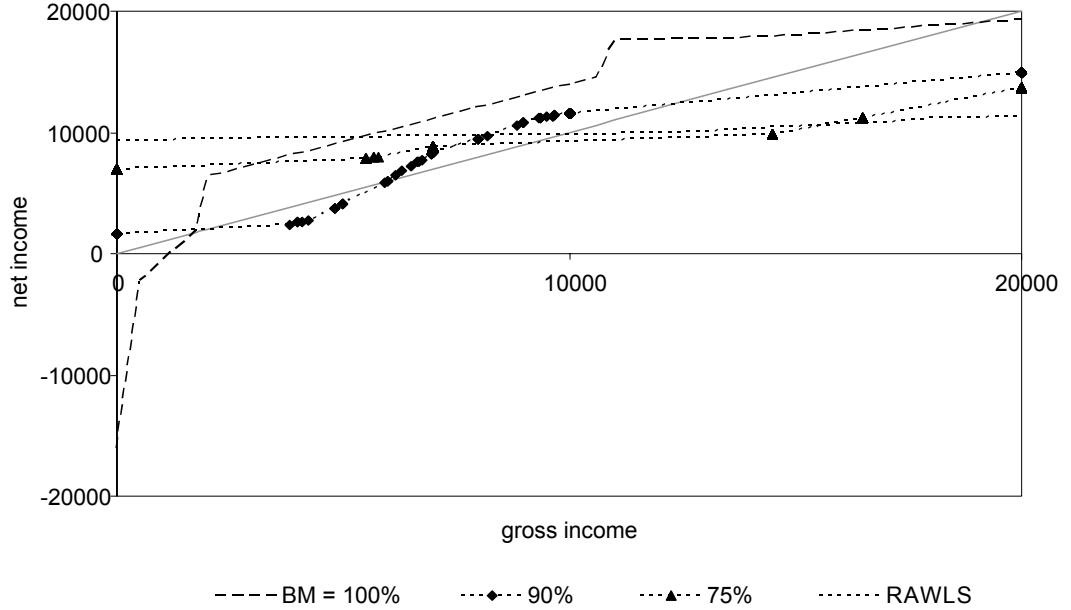


Figure 6: Measuring the impact of constraints.

It clearly shows that taking choice constraints into account dramatically alters our fair tax-benefit scheme. The 90%-case installs a basic income equal to €1631, but it is still followed by a small regressive region (around €5000 - €10000), whereas the 75%-case already leads to a progressive tax-benefit scheme with a basic income equal to €6931 and rather high marginal taxes for the bottom incomes.

6 Conclusion

Given the increased importance many governments attach to "making work pay" policies, we examine whether subsidizing low earners is optimal according to a specific "fair" social ordering. Fairness considerations are kept simple in this paper: we want to compensate individuals for differences in productive skills, but we keep them responsible for their tastes for working.

In a discrete Stiglitz (1982, 1987) economy with four types, optimal taxes crucially depend on the low skill level. Taxes should always be regressive for the bottom incomes—to improve the situation of the worst-off, the hard-working "deserving" poor—unless the low-skilled have zero skills. Although it is reasonable to assume that all individuals with a capacity for work have strictly positive skills, labour market frictions—such as minimum wage laws and job rationing—may nullify their skills. As a consequence, "making work pay" schemes are optimal, only if individuals are unconstrained when making labour choices.

Our simulation results, calibrated on a sample of Belgian singles, illustrate the crucial issue again: are individuals and, more specifically, the unemployed, truly responsible for their labour choices? If all unemployment is voluntary, "making work pay" schemes are optimal. However, this assumption is hardly plausible. If only a small number of

unemployed have similar tastes for working as the employed, but cannot work due to constraints (beyond their responsibility), the optimal tax-benefit scheme changes drastically. More precisely, it moves quickly from a low basic income and high subsidies for the bottom incomes towards a sizeable basic income combined with high marginal taxes for the low earners. The crucial question —the percentage of the unemployed who are willing, but unable to work— is ultimately an empirical issue. Meanwhile, since conservative low estimates change our results rapidly towards a high basic income and high marginal taxes for the bottom incomes, we believe that the latter, rather than making work pay schemes, should be given the benefit of the doubt.

Appendix A: Proofs

Proof of proposition 1

If a rule $f : \mathcal{E} \rightarrow \mathcal{R}$ satisfies Pareto, Anonymity, Compensation and Well-being Independence, then, for each $e \in \mathcal{E}$ and for all allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$: $\min \mathbf{W}(\mathbf{z}, e) > \min \mathbf{W}(\mathbf{z}', e) \Rightarrow \mathbf{z} P_e \mathbf{z}'$.

Proof.

First, we show, in three steps, that Pareto Indifference (the first part of the Pareto axiom) and Well-being Independence for f are equivalent with Neutrality for f :

NEUTRALITY: For all economies $e = (\mathbf{s}, \mathbf{U})$, $e' = (\mathbf{s}, \mathbf{U}') \in \mathcal{E}$ and for all allocations $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathcal{Z}$: If $\mathbf{W}(\mathbf{a}, e) = \mathbf{W}(\mathbf{b}, e')$ and $\mathbf{W}(\mathbf{c}, e) = \mathbf{W}(\mathbf{d}, e')$, then $\mathbf{a} R_e \mathbf{c} \Leftrightarrow \mathbf{b} R_{e'} \mathbf{d}$.

1. If f satisfies Neutrality then f also satisfies Pareto Indifference. Suppose the antecedent of Pareto Indifference is true for a certain economy $e \in \mathcal{E}$ and two allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$, i.e., $U_i(z_i) = U_i(z'_i)$ for all $i \in N$. As such, z_i lies on the same indifference curve as z'_i for all individuals and, by definition of our well-being concept, $\mathbf{W}(\mathbf{z}, e) = \mathbf{W}(\mathbf{z}', e)$. Let $e' = e$ and define allocations $\mathbf{a} = \mathbf{d} = \mathbf{z}$ and $\mathbf{b} = \mathbf{c} = \mathbf{z}'$. As a consequence, $\mathbf{W}(\mathbf{a}, e) = \mathbf{W}(\mathbf{b}, e')$ and $\mathbf{W}(\mathbf{c}, e) = \mathbf{W}(\mathbf{d}, e')$ are true by construction. Using Neutrality, we get $\mathbf{a} R_e \mathbf{c} \Leftrightarrow \mathbf{b} R_{e'} \mathbf{d}$, or equivalently, $\mathbf{z} R_e \mathbf{z}' \Leftrightarrow \mathbf{z}' R_e \mathbf{z}$ (\bullet). Because of completeness of R_e , we must have either $\mathbf{z} R_e \mathbf{z}'$ (and also $\mathbf{z}' R_e \mathbf{z}$ via (\bullet)) or $\mathbf{z}' R_e \mathbf{z}$ (and also $\mathbf{z} R_e \mathbf{z}'$ via (\bullet)). Both cases, lead to $\mathbf{z} I_e \mathbf{z}'$ establishing Pareto Indifference.

2. If f satisfies Neutrality then f also satisfies Well-being Independence. Suppose the antecedent of Well-being Independence is true, i.e., there exist two economies $e = (\mathbf{s}, \mathbf{U})$, $e' = (\mathbf{s}, \mathbf{U}') \in \mathcal{E}$ and two allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$ such that $\mathbf{W}(\mathbf{z}, e) = \mathbf{W}(\mathbf{z}, e')$ and $\mathbf{W}(\mathbf{z}', e) = \mathbf{W}(\mathbf{z}', e')$. Simply choose $\mathbf{a} = \mathbf{b} = \mathbf{z}$ and $\mathbf{c} = \mathbf{d} = \mathbf{z}'$ such that $\mathbf{W}(\mathbf{a}, e) = \mathbf{W}(\mathbf{b}, e')$ and $\mathbf{W}(\mathbf{c}, e) = \mathbf{W}(\mathbf{d}, e')$ holds. Using Neutrality, we get $\mathbf{a} R_e \mathbf{c} \Leftrightarrow \mathbf{b} R_{e'} \mathbf{d}$, or equivalently, $\mathbf{z} R_e \mathbf{z}' \Leftrightarrow \mathbf{z} R_{e'} \mathbf{z}'$ establishing Well-being Independence.

3. If f satisfies Pareto Indifference and Well-being Independence then f also satisfies Neutrality. Suppose the antecedent of Neutrality holds, i.e., there exist two economies $e = (\mathbf{s}, \mathbf{U})$ and $e' = (\mathbf{s}, \mathbf{U}') \in \mathcal{E}$ and four allocations $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathcal{Z}$ such that $\mathbf{W}(\mathbf{a}, e) = \mathbf{W}(\mathbf{b}, e')$ and $\mathbf{W}(\mathbf{c}, e) = \mathbf{W}(\mathbf{d}, e')$. Let us focus on an arbitrary individual $i \in N$. Because $\mathbf{W}_i(\mathbf{a}, e) = \mathbf{W}_i(\mathbf{b}, e')$, the indifference curve of U_i through a_i and U'_i through b_i are tangent to the same “shared resources” opportunity set defined by \mathbf{s} . Given $U_i, U'_i \in \mathcal{U}$, both indifference curves must cross at least once in $\mathbb{R} \times [0, 1]$. Choose a bundle α_i where both cross. Repeating this construction of α_i for all individuals, we get an allocation $\boldsymbol{\alpha} \in \mathcal{Z}$ such that $\mathbf{W}(\boldsymbol{\alpha}, e) = \mathbf{W}(\mathbf{a}, e) = \mathbf{W}(\mathbf{b}, e') = \mathbf{W}(\boldsymbol{\alpha}, e')$. In the

same way, define an allocation $\beta \in \mathcal{Z}$ such that $\mathbf{W}(\beta, e) = \mathbf{W}(\mathbf{c}, e) = \mathbf{W}(\mathbf{d}, e') = \mathbf{W}(\beta, e')$. Using Pareto Indifference and transitivity of R_e and $R_{e'}$, we get:

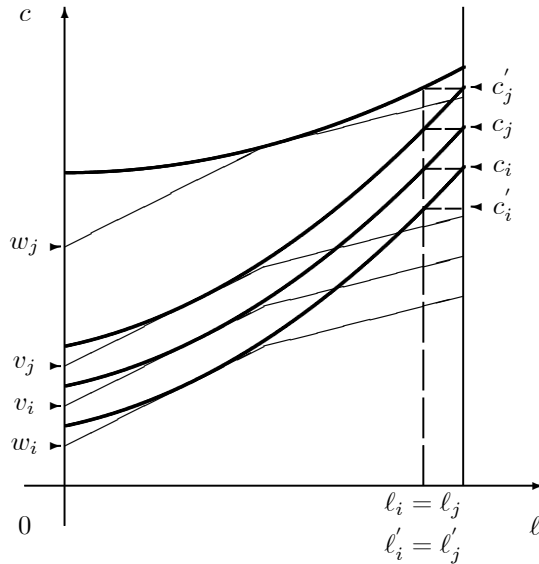
$$(\bullet) \quad \mathbf{a}R_e\mathbf{c} \Leftrightarrow \alpha R_e\beta \text{ and } \alpha R_{e'}\beta \Leftrightarrow \mathbf{b}R_{e'}\mathbf{d}.$$

Using Well-being Independence, we get $\alpha R_e\beta \Leftrightarrow \alpha R_{e'}\beta$. Together with (\bullet) , we get $\mathbf{a}R_e\mathbf{c} \Leftrightarrow \mathbf{b}R_{e'}\mathbf{d}$, establishing Neutrality.

Second, we show that Neutrality is equivalent with welfarism, i.e., a rule f satisfies Neutrality if and only if there exists a unique ordering R^* defined over \mathbb{R}^n , such that, for each economy $e \in \mathcal{E}$ and for all allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$ we have $\mathbf{z}R_e\mathbf{z}'$ if and only if $\mathbf{W}(\mathbf{z}, e) R^* \mathbf{W}(\mathbf{z}', e)$. Here welfarism has to be interpreted as follows: only well-being levels (rather than utility levels) matter to rank two allocations (given a fixed size n of the population and a fixed skill vector \mathbf{s}). Welfarism is a well-known result in the social choice literature (see Bossert and Weymark (2004), theorem 2, for a proof). It suffices to notice that our set-up is sufficiently rich to obtain welfarism: for any two well-being vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, there exist two allocations $\mathbf{z}, \mathbf{z}' \in \mathcal{Z}$ and an economy $e \in \mathcal{E}$ such that $\mathbf{W}(\mathbf{z}, e) = \mathbf{v}$ and $\mathbf{W}(\mathbf{z}', e) = \mathbf{w}$.

Third, the unique ordering R^* inherits certain properties from f : R^* must satisfy weak Pareto (if $v_i > w_i$ for all $i \in N$, then $\mathbf{v}P^*\mathbf{w}$) and Anonymity ($\mathbf{v}I^*\pi(\mathbf{v})$ with $\pi : N \rightarrow N$ a permutation of individuals in N). This is straightforward. We prove that R^* must also satisfy

HAMMOND EQUITY: For all well-being vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and for all individuals $i, j \in N$: If (i) $w_i < v_i < v_j < w_j$ and (ii) $v_k = w_k$ for all $k \neq i, j$, then $\mathbf{v}R^*\mathbf{w}$.



Suppose the antecedent of Hammond Equity holds, or there exist two well-being vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and two individuals $i, j \in N$ such that $w_i < v_i < v_j < w_j$ and $v_k = w_k$ for all $k \neq i, j$ hold. The figure illustrates how it is possible to construct bundles z_i, z_j and

z'_i, z'_j and a utility function $U_i = U_j \in \mathcal{U}$ such that $\mathbf{W}_i(\mathbf{z}, e) = v_i$, $\mathbf{W}_j(\mathbf{z}, e) = v_j$ and $\mathbf{W}_i(\mathbf{z}', e) = w_i$, $\mathbf{W}_j(\mathbf{z}', e) = w_j$ and the antecedents of the Compensation principle are satisfied for i, j , i.e., (i) $\ell_i = \ell_j$, $\ell'_i = \ell'_j$ and $U_i = U_j$ (ii) $\exists \delta > 0$ such that $c'_i = c_i + \delta < c_j - \delta = c'_j$. The bundles z_i, z_j and z'_i, z'_j can be extended with bundles $z_k = z'_k$ for the other individuals $k \neq i, j$ to obtain allocations \mathbf{z} and \mathbf{z}' , such that $\mathbf{W}_k(\mathbf{z}, e) = v_k = w_k = \mathbf{W}_k(\mathbf{z}', e)$ holds for all $k \neq i, j$. Using Compensation, we must have $\mathbf{z} R_e \mathbf{z}'$ and thus, via welfarism, also $\mathbf{v} R^* \mathbf{w}$ must hold.

Finally, Given the axioms for R^* , Tungodden (2000, theorem 1) shows that $\min \mathbf{v} > \min \mathbf{w}$ implies $\mathbf{v} P^* \mathbf{w}$, for any vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, which, given welfarism, completes our proof.

Proof of propositions 2 and 3

To prove propositions 2 and 3, we need two “tricks” and two lemmas. We start with the tricks.

Consider an implementable and feasible allocation $\mathbf{x} \in \mathcal{X}$ as in figure A1. The bundles x_{LL} and x_{HH} lie somewhere in the left and right shaded zone, respectively, to satisfy the incentive constraints. The bundle x° is constructed to satisfy $V_{HL}(x^\circ) = V_{HL}(x_{HL})$ and $MRS_{Y_{HL}}(x^\circ) = 1$.

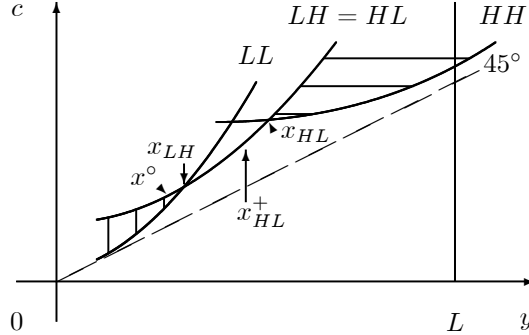


Figure A1: the allocations \mathbf{x} and \mathbf{x}^+ illustrating trick 1.

Now, consider the allocation $\mathbf{x}^+ \in \mathcal{X}$ with $x_{st}^+ = x_{st}$ for all types $st \neq HL$ and x_{HL}^+ is constructed by moving x_{HL} on his indifference curve towards the bundle x° . It is clear that the allocation \mathbf{x}^+ is implementable. Furthermore, given the preference technology defined by \mathcal{U} , we have $y_{HL}^+ - c_{HL}^+ > y_{HL}^* - c_{HL}^*$. Thus, the allocation \mathbf{x}^+ is also feasible, with $\sum_{st \in ST} n_{st} (y_{st}^+ - c_{st}^+) - \sum_{st \in ST} n_{st} (y_{st}^* - c_{st}^*) = m > 0$. The amount of money m can now be freely redistributed to the net income of all types (while still satisfying all incentive constraints) resulting in a weak Pareto improvement and thus also an improvement according to the government's program (*). More generally, we obtain:

TRICK 1: Consider an implementable and feasible allocation $\mathbf{x} \in \mathcal{X}$ and a type st whose bundle x_{st} can be moved along his indifference curve (i) without violating incentive constraints and (ii) making an amount of money m free for redistribution. The allocation \mathbf{x} cannot be optimal according to program (*), because everyone can be made strictly better-off (by redistributing the amount of money m to the net incomes of all types), without violating incentive constraints.

To illustrate the second trick, consider an implementable and feasible allocation $\mathbf{x} \in \mathcal{X}$ as in figure A2.

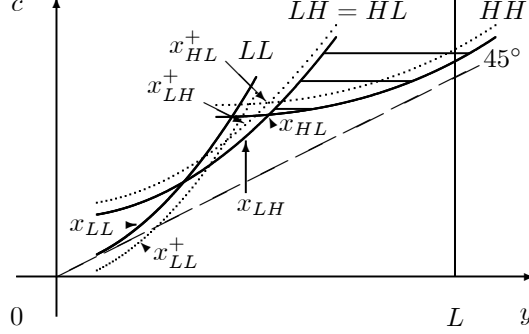


Figure A2: the allocation \mathbf{x} and \mathbf{x}^+ illustrating trick 2.

Again, the bundle x_{HH} lies somewhere in the right shaded zone to satisfy the incentive constraints. Now it is possible to construct a feasible and implementable allocation $\mathbf{x}^+ \in \mathcal{X}$, transferring in \mathbf{x} some net income from type LL to the other types LH , HL and HH . Whether or not the resulting allocation is better according to program (*), ultimately depends on the well-being levels in society: if LL is strictly better off compared to one of the other types, it is always possible to find an allocation \mathbf{x}^+ which is better according to program (*). We summarize

TRICK 2: Consider an implementable and feasible allocation $\mathbf{x} \in \mathcal{X}$ and one or more types st whose bundle(s) x_{st} can be moved downwards (i) without violating incentive constraints and (ii) making an amount of money $m > 0$ free for redistribution to the other types. The allocation \mathbf{x} cannot be optimal according to program (*), if all donor type(s) st were strictly better off in \mathbf{x} compared to (one of) the other types.

Besides two tricks, we need two lemmas. The first lemma tells us that the program (*) can, loosely speaking, focus on the lower-skilled, because they are always worse-off in terms of well-being, more precisely:

Lemma 1. Consider two types with the same taste for working $t \in T$ (and thus the same utility function $U_t \in \mathcal{U}$), but different skills $0 \leq L < H$. In an implementable allocation $\mathbf{x} \in \mathcal{X}$, with $V_{Ht}(x_{Ht}) \geq V_{Ht}(x_{Lt})$ (resp. $V_{Ht}(x_{Ht}) > V_{Ht}(x_{Lt})$) the lower-skilled type Lt is always worse off (resp. strictly worse off) compared to the higher-skilled type Ht , i.e., $\mathbf{W}_{Ht}(\mathbf{x}) \geq \mathbf{W}_{Lt}(\mathbf{x})$ (resp. $\mathbf{W}_{Ht}(\mathbf{x}) > \mathbf{W}_{Lt}(\mathbf{x})$).

Proof. Consider two types with the same taste for working $t \in T$. We prove the case where skills satisfy $0 < L < T$ and $V_{Ht}(x_{Ht}) \geq V_{Ht}(x_{Lt})$; the other cases are analogous. Call (c_{Lt}, y_{Lt}) and (c_{Ht}, y_{Ht}) their bundles. Individuals with the same taste t have the same utility functions U_t and thus also the same indifference curves and therefore the same well-being level for bundles on the same indifference curve. Because our well-being measure is ordinaly equivalent with utility, measured by U_t , it suffices to show that $U_t(c_{Lt}, \frac{y_{Lt}}{L}) \leq U_t(c_{Ht}, \frac{y_{Ht}}{H})$. Suppose not, i.e., suppose (i) $U_t(c_{Lt}, \frac{y_{Lt}}{L}) > U_t(c_{Ht}, \frac{y_{Ht}}{H})$.

Because $V_{Ht}(x_{Ht}) \geq V_{Ht}(x_{Lt})$ we get, by definition of V_{Ht} , that (ii) $U_t(c_{Ht}, \frac{y_{Ht}}{H}) \geq U_t(c_{Lt}, \frac{y_{Lt}}{L})$. Combining (i) and (ii), we obtain $U_t(c_{Lt}, \frac{y_{Lt}}{L}) > U_t(c_{Lt}, \frac{y_{Lt}}{H})$, a contradiction given $U_t \in \mathcal{U}$ and $0 < L < H$. \square

Lemma 2 tells us that it cannot be optimal —according to the government's program (*)— to treat the indistinguishable middle types LH and HL differently in case $y_{HL}^* \leq L$. Otherwise (if $y_{HL}^* > L$) it might be optimal to treat them differently, but only under certain conditions:

Lemma 2. *Consider a four type economy with skills $0 < L < H$ and tastes represented by utility functions $U_L, U_H \in \mathcal{U}$, which satisfy single-crossingness and indistinguishable middle type. Consider a government who optimizes the program defined by (*). In an optimal allocation $\mathbf{x}^* \in \mathcal{X}$, we must have:*

- (a). $x_{LH}^* = x_{HL}^*$, if $y_{HL}^* \leq L$, or else,
- (b). $V_{HL}(x_{LH}^*) = V_{HL}(x_{HL}^*)$, with $y_{LH}^* = L < y_{HL}^*$ and $MRSY_{HL}(x_{HL}^*) \leq 1$.

Proof of part (a). Suppose $y_{HL}^* \leq L$ and $x_{LH}^* \neq x_{HL}^*$. We show that it is always possible to construct another allocation $\mathbf{x} \in \mathcal{X}$, which is feasible, implementable and strictly better than \mathbf{x}^* according to the government's program (*). Because $y_{HL}^* \leq L$, the incentive compatibility constraints $IC_{LH,HL}$ and $IC_{HL,LH}$ require

$$\begin{aligned} V_{HL}(x_{HL}^*) &\geq V_{HL}(x_{LH}^*) \text{ and} \\ V_{LH}(x_{LH}^*) &\geq V_{LH}(x_{HL}^*) \Leftrightarrow V_{HL}(x_{LH}^*) \geq V_{HL}(x_{HL}^*), \end{aligned}$$

where the equivalence \Leftrightarrow is due to indistinguishable middle types. We must have $V_{HL}(x_{HL}^*) = V_{HL}(x_{LH}^*)$, or x_{LH}^* and x_{HL}^* must lie on the same indifference curve.

Given our preference technology \mathcal{U} , there are only two cases for $x_{LH}^* \neq x_{HL}^*$. Assume $x_{LH}^* < x_{HL}^*$, i.e., $c_{LH}^* < c_{HL}^*$ and $y_{LH}^* < y_{HL}^*$; for the other case $x_{LH}^* > x_{HL}^*$ simply switch subscripts HL and LH in the sequel. Define a bundle $x^\circ = (c^\circ, y^\circ)$ in $\mathbb{R} \times [0, L]$ such that x° also lies on the same indifference curve through x_{LH}^* and x_{HL}^* , i.e., $V_{HL}(x^\circ) = V_{HL}(x_{HL}^*)$, and choose (i) $y^\circ = 0$, if $MRSY_{HL} \geq 1$ everywhere in $\mathbb{R} \times [0, L]$, (ii) $y^\circ = L$, if $MRSY_{HL} \leq 1$ everywhere in $\mathbb{R} \times [0, L]$, or else (iii) choose x° such that $MRSY_{HL}(x^\circ) = 1$. Each case leads to any of the following three cases: either (α) $y^\circ \leq y_{LH}^* < y_{HL}^*$, or (β) $y_{LH}^* < y^\circ < y_{HL}^*$ or (γ) $y_{LH}^* < y_{HL}^* \leq y^\circ$. In each of the three cases (α), (β) and (γ), it is possible to use TRICK 1, by moving either x_{HL}^* to the left on his indifference curve (in case (α) and (β)) or moving x_{LH}^* to the right on his indifference curve (in case (γ)), contradicting that \mathbf{x}^* was optimal.

Proof of part (b). Suppose $y_{HL}^* > L$. We show that $x_{LH}^* < x_{HL}^*$, with $y_{LH}^* = L$ and $MRSY_{HL}(x_{HL}^*) \leq 1$, must hold. Recall that, in case $y_{HL}^* > L$, the incentive constraint $IC_{LH,HL}$ does not exist, because type HL 's bundle is not attainable for LH .

We first show that the incentive constraint $IC_{HL,LH}$ must bind, i.e., $V_{HL}(x_{HL}^*) = V_{HL}(x_{LH}^*)$. Suppose not, i.e., $V_{HL}(x_{HL}^*) > V_{HL}(x_{LH}^*)$. Single-crossingness ensures that $V_{HL}(x_{HL}^*) > V_{HL}(x_{LL}^*)$ and thus LL is strictly worse-off compared to HL (lemma 1); for the same reason, LH is strictly worse off compared to HH . Now, it is possible to use TRICK 2, transferring from type HL (and possibly HH as well, if $IC_{HL,HH}$ binds) to both other types LL and LH , which must improve the lowest well-being, contradicting that \mathbf{x}^* was optimal according to program (*).

Now, we are back in the same situation as in part (a), because both x_{LH}^* and x_{HL}^* , with $x_{LH}^* < x_{HL}^*$, lie on the same indifference curve (of type HL), i.e., $V_{HL}(x_{LH}^*) = V_{HL}(x_{HL}^*)$, but here $y_{LH}^* \leq L < y_{HL}^*$. Now proceed as in part (a). Define the bundle $x^\circ = (c^\circ, y^\circ)$ in $\mathbb{R} \times [0, H]$ such that x° also lies on the same indifference curve through x_{LH}^* and x_{HL}^* , i.e., $V_{HL}(x^\circ) = V_{HL}(x_{HL}^*)$, and choose (i) $y^\circ = 0$, if $MRSY_{HL} \geq 1$ everywhere in $\mathbb{R} \times [0, H]$, (ii) $y^\circ = H$, if $MRSY_{HL} \leq 1$ everywhere in $\mathbb{R} \times [0, H]$, or else (iii) choose x° such that $MRSY_{HL}(x^\circ) = 1$. Now, $y^\circ < y_{HL}^*$ is not possible (otherwise we can use TRICK 1, moving x_{HL}^* to the left on his indifference curve); thus $MRSY_{HL}(x_{HL}^*) \leq 1$. As a consequence $y^\circ \geq y_{HL}^*$ must hold. Now, $y_{LH}^* < L$ is not possible (because then $y_{LH}^* < L < y^\circ$ and using trick 1 again, we could move x_{LH}^* to the right on his indifference curve). Thus $y_{LH}^* = L$, which completes the proof. \square

We are ready to prove propositions 2 and 3, on the basis of lemmas 1 and 2 and tricks 1 and 2.

Proof of proposition 2

Consider a four type economy with skills $0 < L < H$ and tastes represented by utility functions $U_L, U_H \in \mathcal{U}$, which satisfy single-crossingness and indistinguishable middle type. Consider a government who optimizes the program defined by (*). In an optimal allocation $\mathbf{x}^* \in \mathcal{X}$, we must have $y_{LL}^* - c_{LL}^* \geq y_{LH}^* - c_{LH}^*$.

Our proof consists of two parts, depending on whether (a) $y_{HL}^* \leq L$ in the optimum \mathbf{x}^* , or (b) $y_{HL}^* > L$. Given the definition of \mathcal{X} , one of both cases must hold. We show, for both cases, that $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$ is not possible.

Proof of part (a). Suppose $y_{HL}^* \leq L$ (thus $x_{LH}^* = x_{HL}^*$ via lemma 2) and $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$. We consider four possible cases, depending on whether $IC_{LL,LH}$ and/or $IC_{LH,LL}$ bind, or not. In an optimum \mathbf{x}^* of the program (*), one of these four cases must hold. For all cases, we show that it is possible to construct a strictly better allocation according to the program (*), which also satisfies the feasibility and incentive compatibility constraints.

1. $IC_{LL,LH}$ and $IC_{LH,LL}$ bind. This requires $x_{LL}^* = x_{LH}^*$ which contradicts $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$.

2. $IC_{LL,LH}$ binds, $IC_{LH,LL}$ does not bind. **2a.** If $MRSY_{LL}(x_{LL}^*) < 1$, we could use TRICK 1 moving x_{LL}^* somewhat to the right on his indifference curve. **2b.** We must have $MRSY_{LL}(x_{LL}^*) \geq 1$, from (2a). But, given the preference technology defined by \mathcal{U} , $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$ is not possible, a contradiction.

3. $IC_{LL,LH}$ does not bind, $IC_{LH,LL}$ binds: **3a.** Let us first focus on type LH . If $y_{LH}^* = 0$, then incentive constraints and single-crossingness require $x_{LL}^* = x_{LH}^*$, which violates $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$. So $y_{LH}^* > 0$. If $MRSY_{LH}(x_{LH}^*) > 1$, we can use TRICK 1 again, by moving both $x_{LH}^* = x_{HL}^*$ to the left on their (common) indifference curve. So $MRSY_{LH}(x_{LH}^*) \leq 1$ must hold. **3b.** We focus now on type LL . We must have either (i) $y_{LL}^* = 0$ or (ii) $y_{LL}^* > 0$. In case (ii), we have $MRSY_{LH}(x_{LH}^*) \leq 1$ (otherwise we can use TRICK 1, moving x_{LH}^* somewhat to the left on his indifference curve). **3c.** Due to lemma 1, either type LH or LL has the minimal well-being. Figure A3 illustrates (3a), (3b(ii)) and $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$; type HH 's bundle is somewhere in the shaded zone. It is easy to verify that type LH is always strictly worse-off compared to type

LL , irrespective of the proportion of high-skilled $\frac{n_{HL}+n_{HH}}{n}$ (which defines the kink in the budget set where the slope changes from $\frac{H}{L} > 1$ to 1) and irrespective of whether (3b(i)) or (3b(ii)) applies.

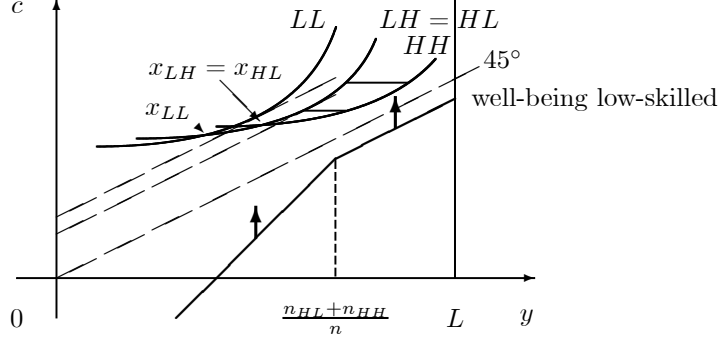


Figure A3: type LL is strictly better off compared to type LH .

Since $IC_{LL,LH}$ does not bind, it is always possible to use TRICK 2 transferring a small amount of money from LL to the other types LH , HL and HH , improving the minimal well-being in society, a contradiction.

4. $IC_{LL,LH}$ and $IC_{LH,LL}$ do not bind. Using TRICK 1, it can be verified that only the following cases are possible: (i) $y_{LL}^* = 0 < y_{LH}^*$, $MRSY_{LL}(x_{LL}^*) \geq 1$ and $MRSY_{LH}(x_{LH}^*) \leq 1$, or (ii) $0 < y_{LL}^* < y_{LH}^*$, $MRSY_{LL}(x_{LL}^*) = 1$ and $MRSY_{LH}(x_{LH}^*) \leq 1$. We are back in the same situation as in (3). In both cases (i) and (ii) and, given $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$, type LH is strictly worse-off compared to LL , irrespective of the proportion of high-skilled. Here again, TRICK 2 can be used to obtain a contradiction.

Proof of part (b). Suppose $y_{HL}^* > L$ (thus $V_{HL}(x_{LH}^*) = V_{HL}(x_{HL}^*)$, with $y_{LH}^* = L < y_{HL}^*$ and $MRSY_{HL}(x_{HL}^*) \leq 1$ via lemma 2) and $y_{LL}^* - c_{LL}^* < y_{LH}^* - c_{LH}^*$. It is again possible to consider four cases, depending on whether $IC_{LL,LH}$ and/or $IC_{LH,LL}$ bind or not, and to show, for each case, a contradiction. Actually, the proof is completely analogous as in steps 1-4 of part (a) and therefore omitted.

Proof of proposition 3

Consider a four type economy with skills $L = 0 < H$ and tastes represented by utility functions $U_L, U_H \in \mathcal{U}$, which satisfy single-crossingness and indistinguishable middle type. Consider a government who optimizes the program defined by (*). In an optimal allocation $\mathbf{x}^* \in \mathcal{X}$ we must have $y_{LL}^* - c_{LL}^* = y_{LH}^* - c_{LH}^* \leq y_{HL}^* - c_{HL}^*$.

Proof. Suppose $y_{LL}^* - c_{LL}^* = y_{LH}^* - c_{LH}^* > y_{HL}^* - c_{HL}^*$ holds. **1.** Because $L = 0$ and $\mathbf{x}^* \in \mathcal{X}$, we must have $y_{LL}^* = y_{LH}^* = 0$ and, given the incentive constraints, also $c_{LL}^* = c_{LH}^*$ must hold. **2.** Due to lemma 1, the lowest well-being is either LH or LL , thus, given (1), we must maximize the basic income, i.e., maximize $c_{LL}^* = c_{LH}^*$. **3.** $y_{HL}^* = 0$ is excluded, otherwise we must have $c_{LH}^* = c_{HL}^*$ (due to incentive constraints) and $y_{LH}^* - c_{LH}^* > y_{HL}^* - c_{HL}^*$ would be violated. **4.** So, $y_{HL}^* > 0$ holds from (3). Now, x_{LH}^* and x_{HL}^* must lie on the same indifference curve, or $V_{HL}(x_{HL}^*) = V_{HL}(x_{LH}^*)$. Otherwise (see the proof of lemma 2, part (b)) it would be possible to improve the situation of the worst-off types LL and LH , at the cost of the better-off types HL and

HH (on the basis of TRICK 2). **5.** If $MRSY_{HL}(x_{HL}^*) > 1$ at $y_{HL}^* > 0$, we can use TRICK 1, moving x_{HL}^* to the left on his indifference curve. **6.** To summarize, we must have $MRSY_{HL}(x_{HL}^*) \leq 1$ and $y_{HL}^* > 0$ while $V_{HL}(x_{HL}^*) = V_{HL}(x_{LH}^*)$ and $y_{LH}^* = 0$. But this contradicts $y_{LH}^* - c_{LH}^* > y_{HL}^* - c_{HL}^*$, given our preference technology defined by \mathcal{U} .

Appendix B: The Belgian tax system for singles

pre-tax income y	marginal tax rate (in %)
$< \text{€}5032$	0
$\text{€}5033 - \text{€}6272$	25
$\text{€}6273 - \text{€}8304$	30
$\text{€}8305 - \text{€}1849$	40
$\text{€}11850 - \text{€}27268$	45
$\text{€}27269 - \text{€}40902$	50
$\text{€}40903 - \text{€}59990$	52.5
$> \text{€}59990$	55

Appendix C: Imputation via a sample selection model

First, we present the variables used for imputing gross hourly wages σ and net hourly benefits b_h ; afterwards, we show the estimates for both sample selection models.

Imputing gross hourly wages σ

In the wage equation, the independent variables are:

- age and its square (*age*, *agesq*)
- educational dummies indicating the highest achieved education level of the individual, starting from primary education (base case), lower secondary education (*dumeduc2*), higher secondary education (*dumeduc3*), higher education short type (*dumeduc4*), higher education long type (*dumeduc5*)
- a gender dummy (*sex*) taking the value of 1 for females

In the selection equation, the independent variables are:

- physical health dummies indicating the general health situation of the individual, ranging from very good (base case), good (*dumhealth2*), reasonable (*dumhealth3*), to bad (*dumhealth4*)
- mental health dummies indicating how often the individual feels depressed, ranging from never (base case), seldom (*dumdepri2*), at times (*dumdepri3*), regularly (*dumdepri4*), to frequently (*dumdepri5*); and how often the individual longs for death, ranging from never (base case), seldom (*dumdeath2*), at times (*dumdeath3*), regularly (*dumdeath4*), to frequently (*dumdeath5*)
- smoking dummies indicating smoking behaviour, ranging from never (base case), occasionally (*dumsmoke2*), to daily (*dumsmoke3*)
- care dummies indicating whether the individual has to take care for his children (*child*) taking the value of 1 if affirmative; and/or has to take care of others (*depperson*) taking the value of 1 if affirmative
- the independent variables of the wage equation

Imputing net hourly benefits b_h

The dependent variable is the marginal benefit per hour $b_h = \frac{b}{2925}$, with b the net yearly benefit. The independent variables are identical to those in the Heckman selection model imputing σ . In addition, we add in both the benefit and the selection equation civil status dummies, indicating whether the individual is divorced (*divorce*), taking the value of 1 if affirmative; widowed (*widow*), taking the value of 1 if affirmative; living together (*cohabit*), taking the value of 1 if affirmative.

Number of observations = 644, from which 136 censored and 508 uncensored. Wald's $\chi^2(7) = 265.68$ with $\Pr > \chi^2(7) = 0.00$ and likelihood ratio test of independent wage and selection equations results in $\chi^2(1) = 4.66$ with $\Pr > \chi^2(1) = 0.031$.

wage equation	Coefficient	Standard Error	$\Pr > z $
age	0.497	0.123	0.000
agesq	-0.003	0.002	0.075
dumeduc2	1.364	1.141	0.232
dumeduc3	2.834	1.103	0.010
dumeduc4	4.508	1.160	0.000
dumeduc5	5.617	1.149	0.000
sex	-0.917	0.383	0.017
cons	-2.680	2.568	0.297
selection equation	Coefficient	Standard Error	$\Pr > z $
dumhealth2	-0.576	0.176	0.001
dumhealth3	-0.686	0.227	0.002
dumhealth4	-1.221	0.476	0.010
dumdepri2	0.014	0.186	0.938
dumdepri3	-0.243	0.185	0.188
dumdepri4	-0.486	0.249	0.051
dumdepri5	-0.799	0.342	0.019
dumdeath2	0.374	0.173	0.031
dumdeath3	0.113	0.196	0.563
dumdeath4	0.160	0.297	0.591
dumdeath5	-0.657	0.352	0.062
dumsmoke2	-0.041	0.245	0.868
dumsmoke3	-0.219	0.138	0.114
child	-0.355	0.153	0.020
depperson	-0.232	0.210	0.269
age	0.179	0.040	0.000
agesq	-0.002	0.001	0.000
dumeduc2	0.587	0.254	0.021
dumeduc3	0.784	0.243	0.001
dumeduc4	1.584	0.305	0.000
dumeduc5	1.685	0.291	0.000
sex	-0.507	0.139	0.000
cons	-2.016	0.728	0.006

Number of observations = 638, from which 480 censored and 158 uncensored. Wald's $\chi^2(10) = 50.51$ with $\Pr > \chi^2(7) = 0.00$ and likelihood ratio test of independent benefit and selection equations results in $\chi^2(1) = 7.09$ with $\Pr > \chi^2(1) = 0.008$.

benefit equation	Coefficient	Standard Error	$\Pr > z $
age	0.115	0.051	0.025
agesq	-0.001	0.001	0.120
dumeduc2	0.361	0.259	0.163
dumeduc3	0.470	0.271	0.083
dumeduc4	0.181	0.358	0.613
dumeduc5	0.575	0.410	0.161
sex	-0.317	0.187	0.091
divorce	0.376	0.197	0.057
widow	-0.504	0.582	0.386
cohabit	-0.035	0.196	0.858
cons	-0.226	0.921	0.806
selection equation	Coefficient	Standard Error	$\Pr > z $
dumhealth2	0.304	0.154	0.048
dumhealth3	0.416	0.203	0.040
dumhealth4	0.384	0.415	0.355
dumdepri2	0.137	0.163	0.400
dumdepri3	0.244	0.166	0.142
dumdepri4	0.470	0.237	0.048
dumdepri5	0.334	0.329	0.311
dumdeath2	-0.464	0.154	0.003
dumdeath3	0.020	0.173	0.909
dumdeath4	0.369	0.254	0.145
dumdeath5	0.889	0.339	0.009
dumsmoke2	0.114	0.214	0.596
dumsmoke3	0.228	0.124	0.065
child	0.304	0.143	0.033
depperson	0.158	0.192	0.410
age	-0.108	0.040	0.007
agesq	0.001	0.001	0.006
dumeduc2	-0.376	0.242	0.120
dumeduc3	-0.707	0.232	0.002
dumeduc4	-1.049	0.265	0.000
dumeduc5	-1.428	0.270	0.000
sex	0.373	0.128	0.004
divorce	0.027	0.162	0.868
widow	0.036	0.490	0.941
cohabit	-0.329	0.144	0.022
cons	1.247	0.717	0.082

Appendix D: Some descriptive statistics

We report the number of observations (n), the minimum (min), the 25th percentile (p25), the median (p50), the 75th percentile (p75) and the maximum (max) for the following variables: age, normalized labour ℓ , observed gross hourly wages σ , imputed gross hourly wages $\hat{\sigma}$, observed and imputed gross hourly wages $(\sigma, \hat{\sigma})$, observed net hourly benefits b_h , imputed net hourly benefits \hat{b}_h and observed and imputed net hourly benefits (b_h, \hat{b}_h) .

	n	min	p25	p50	p75	max
age	621	16	25	33	42	70
ℓ	621	0.000	0.231	0.554	0.615	1.000
σ	502	1.495	9.585	11.864	15.700	38.271
$\hat{\sigma}$	119	1.262	4.796	6.665	9.421	19.214
$\sigma, \hat{\sigma}$	621	1.262	8.594	11.065	14.404	38.271
b_h	143	0.166	0.915	1.849	3.051	3.661
\hat{b}_h	478	0.002	0.072	0.214	0.510	2.089
b_h, \hat{b}_h	621	0.002	0.115	0.345	0.928	3.661

References

- [1] Bernardi, L. and Profeta, P. (2004) *Tax Systems and Tax Reforms in Europe*, New York: Routledge.
- [2] Blundell, R. and MaCurdy, T. (1999) Labor supply: A review of alternative approaches, Chapter 27 in, Ashenfelter, O. and Card, D. (eds.), *Handbook of Labor Economics* 3A, Amsterdam: North Holland.
- [3] Boadway, R., Marchand, M., Pestieau, P. and Racionero, M.D.M. (2002) Optimal redistribution with heterogeneous preferences for leisure, *Journal of Public Economic Theory* 4, 475-498.
- [4] Bossert, W. and Weymark, J. (2004) Utility in social choice, Chapter 20 in, Barbera, S., Hammond, P.J. and Seidl, C. (eds.), *Handbook Of Utility Theory. Volume 2: Extensions*, Dordrecht: Kluwer.
- [5] Diamond, P.A. (1980) Income taxation with fixed hours of work, *Journal of Public Economics* 13, 101-110.
- [6] Diamond, P.A. (1998) Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates, *American Economic Review* 88(1), 83-95.
- [7] Dworkin, R. (1981) What is equality? Part 2: equality of resources, *Philosophical Public Affairs* 10(4), 283-345.
- [8] Ebert, U. (1992) A reexamination of the optimal nonlinear income tax, *Journal of Public Economics* 49(1), 47-73.
- [9] Fleurbaey, M. and Maniquet, F. (1996) Fair allocation with unequal production skills: the no envy approach to compensation, *Mathematical Social Sciences* 32(1), 71-93.
- [10] Fleurbaey, M. and Maniquet, F. (2000) Fair social orderings with unequal production skills, THEMA working paper 2000-17.
- [11] Fleurbaey, M. and Maniquet, F. (2002) Fair income tax, mimeo, Université de Pau.
- [12] Heckman, J. (1993) What has been learnt about labor supply in the past twenty years?, *American Economic Review* 83(2), 116-121.
- [13] Kanbur, R. and Tuomala, M. (1984) Inherent inequality and the optimal graduation of marginal tax rates, *Scandinavian Journal of Economics* 96(2), 275-282.
- [14] Maniquet, F. (1998) An equal right solution to the compensation-responsibility dilemma, *Mathematical Social Sciences* 35(2), 185-202.
- [15] Mirrlees, J.A. (1971) An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38(114), 175-208.
- [16] Pazner, E.A. and Schmeidler, D. (1978) Egalitarian equivalent allocations: a new concept of economics equity, *Quarterly Journal of Economics* 92, 671-687.

- [17] Pearson, M. and Scarpetta, S. (2000) An overview: What do we know about policies to make work pay?, *OECD Economic Studies* 31.
- [18] Piketty, T. (1997) La redistribution fiscale face au chômage, *Revue Française d'Economie* 12(1), 157-201.
- [19] Rawls, J. (1971) *A Theory of Justice*, Cambridge: Harvard University Press.
- [20] Roemer, J., Aaberge, R., Colombino, U., Fritzell, J., Jenkins S.P., Lefranc, A., Marx, I., Page, M., Pommer, E., Ruiz-Castillo, J., San Segundo, M.J., Tranaes, T., Trannoy, A., Wagner, G.G. and Zubiri, I. (2003), To what extent do fiscal regimes equalize opportunities for income acquisition among citizens?, *Journal of Public Economics* 87, 539-565.
- [21] Saez, E. (2001) Using elasticities to derive optimal income tax rates, *Review of Economic Studies* 68, 205-229.
- [22] Saez, E. (2002) Optimal income transfer programs: intensive versus extensive labor supply responses, *Quarterly Journal of Economics* 117(3), 1039-1073.
- [23] Schokkaert, E., Van de gaer, D., Vandenbroucke, F. and Luttens, R.I. (2004), Responsibility-sensitive egalitarianism and optimal linear income taxation, *Mathematical Social Sciences* 48(2), 151-182.
- [24] Seade, J.K. (1977) On the shape of optimal tax schedules, *Journal of Public Economics* 7(2), 203-236.
- [25] Stiglitz, J.E. (1982) Self-selection and Pareto efficient taxation, *Journal of Public Economics* 17, 213-240.
- [26] Stiglitz, J.E. (1987) Pareto Efficient and Optimal Taxation and the New New Welfare Economics, in Auerbach, A.J. and Feldstein, M. (eds.), *Handbook of Public Economics*, volume 2, North-Holland: Elsevier.
- [27] Tungodden, B. (2000) Egalitarianism: is leximin the only option?, *Economics and Philosophy* 16, 229-245.
- [28] Tuomala, M. (1990) *Optimal Income Tax and Redistribution*, Oxford: Clarendon Press.
- [29] Varian, H.R. (1974) Equity, envy and efficiency, *Journal of Economic Theory* 9, 63-91.

